Graphical user interface, application

Description automatically generated

Apuntes Unidad 4

Intervalos de confianza



Shape, arrow

Description automatically generated

**INTERVALOS DE CONFIANZA**

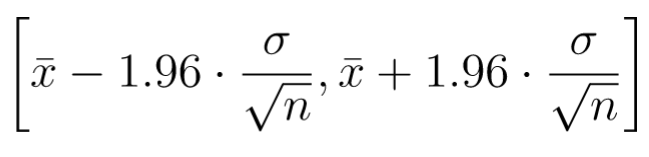
Recordemos que cuando la media poblacional es desconocida no se puede determinar de forma precisa, peso si es posible estimarla a través de un intervalo que contiene a con cierta probabilidad. Para determinar un intervalo de confianza para la media de la población hacemos dos supuestos:

* La distribución de la población sigue una distribución aproximadamente normal.
* El valor de la desviación estándar de la población es conocido.

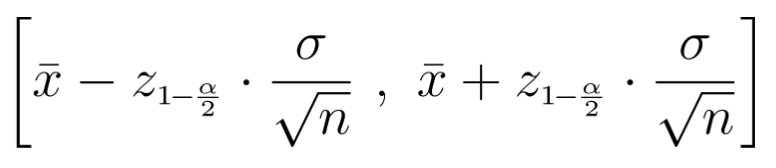
Por ejemplo, fijando la probabilidad de encontrar a en un 95%, tenemos que el intervalo de confianza viene dado por:



Donde corresponde a la media muestral, a la desviación estándar poblacional y al tamaño de la muestra. Ahora bien, es importante recalcar que el intervalo obtenido es aleatorio, pues se genera a partir de la media muestral . El promedio muestral observado, denotado por , corresponde a una realización de la variable aleatoria . Manteniendo el nivel de confianza en 95% y reemplazando por el promedio muestral observado , el intervalo fijo resultante se conoce como intervalo de confianza del 95% para . Este intervalo puede expresarse como:



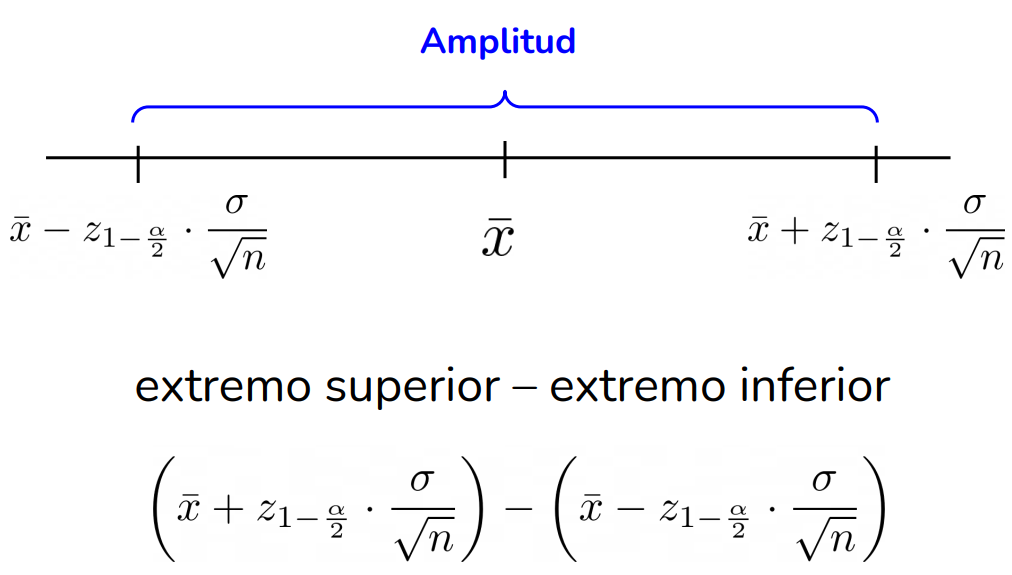
Finalmente, un intervalo de confianza de para una media poblacional , se puede expresar de la siguiente manera:



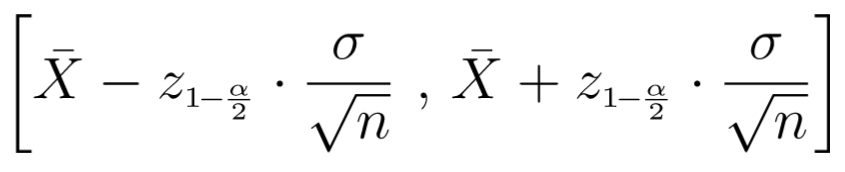
donde corresponde al promedio muestral observado, al tamaño de la muestra extraída desde la población, a la desviación estándar de la población y es el valor de asociado a un nivel de confianza .

**AMPLITUD DE UN INTERVALO DE CONFIANZA**

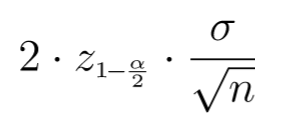
La amplitud de un intervalo de confianza corresponde a la diferencia entre el extremo superior y el extremo inferior del intervalo, como se muestra a continuación:



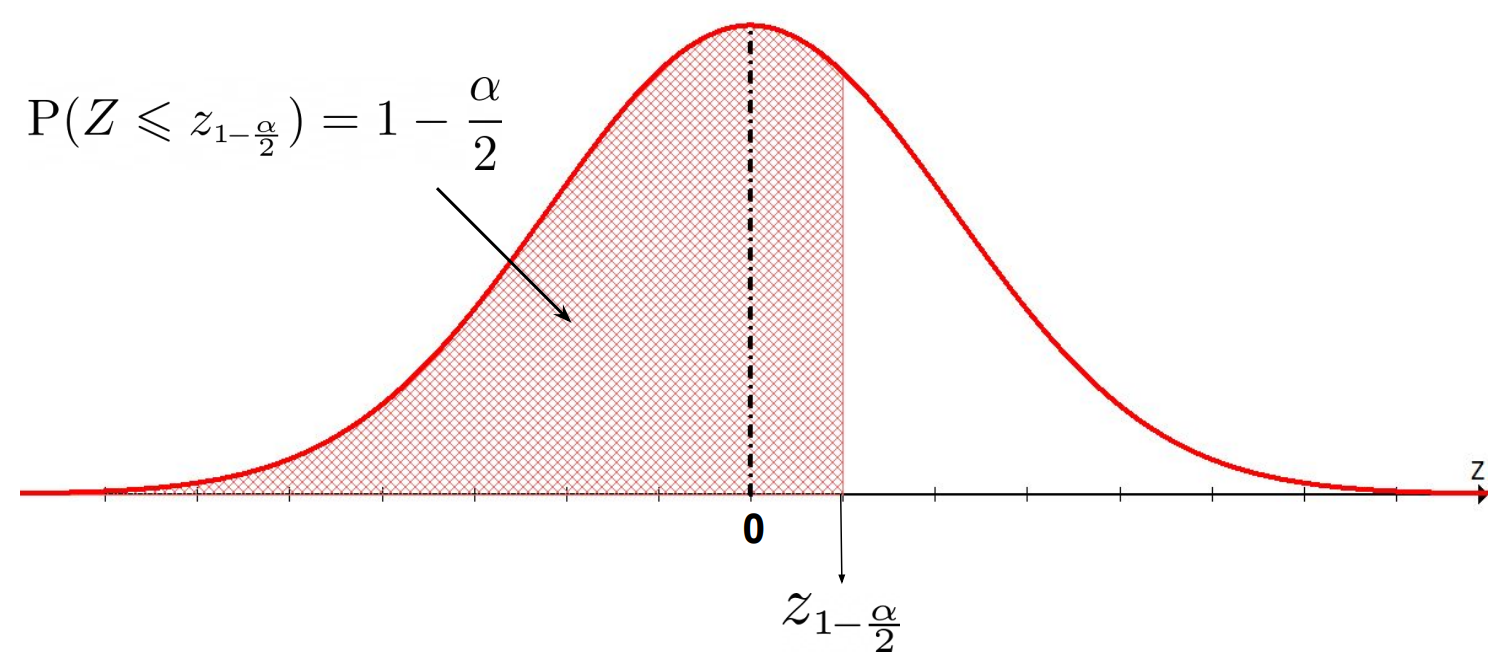
Considerando una población de interés con media desconocida, desviación estándar conocida y suponiendo que seleccionamos una muestra aleatoria de tamaño de esa población y calculamos el promedio muestral , un intervalo que permite estimar con nivel de confianza es:



La amplitud de este intervalo de confianza es igual a:

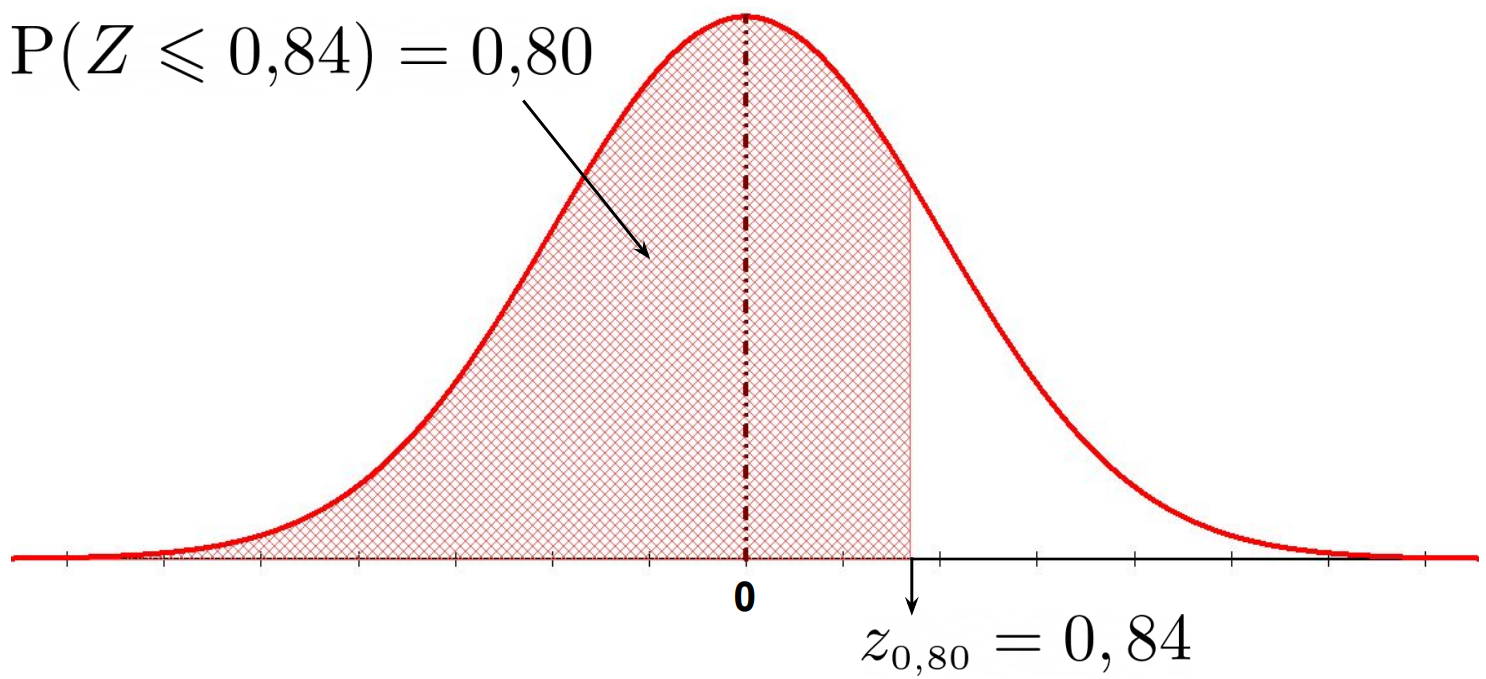
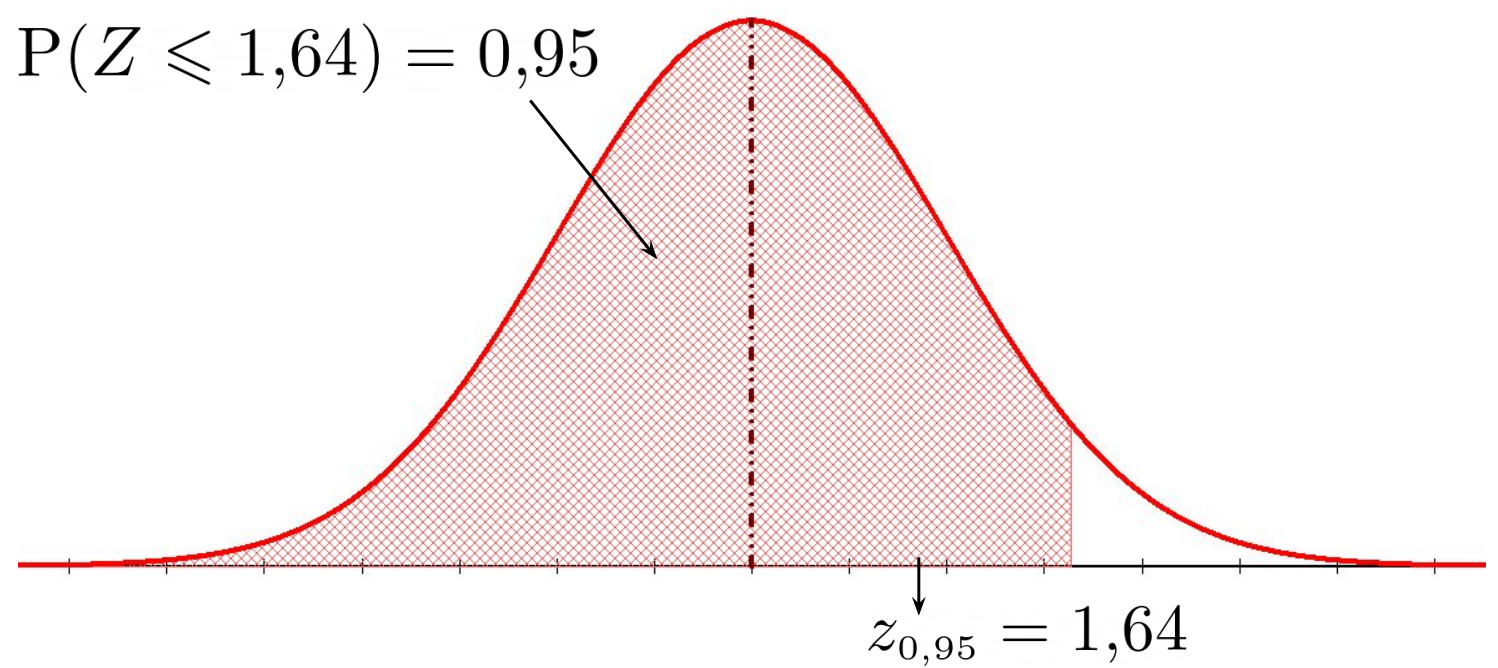


Notemos que cuando los valores de y permanecen constantes, esta amplitud solo cambia al modificar el valor del percentil . Este percentil representa el valor de la variable aleatoria que acumula un área bajo la curva normal estándar igual a , como se muestra en la siguiente figura:



Si cambiamos el nivel de confianza , también cambiará el valor de , ya que ambos están relacionados por medio de , y por tanto también se modificará el valor del percentil .

De hecho, cuando aumenta el nivel de confianza , el valor del percentil también aumenta, ya que el área en rojo debe ser mayor. Los siguientes gráficos ilustran esto:



Como la desviación estándar no cambia (es un parámetro de la población) y dado que la amplitud no depende del promedio muestral , al fijar el nivel de confianza , la amplitud del intervalo de confianza es inversamente proporcional a . Esto se traduce en que, si crece, la amplitud disminuye.

**INTERPRETACIÓN DEL INTERVALO DE CONFIANZA**

Consideremos una población donde se ha observado la variable de interés y se busca estimar la media de esta variable. Para ello, podemos seleccionar una muestra de tamaño de dicha población, calcular el promedio de dicha muestra , y utilizar este para construir un intervalo con nivel de confianza .

Imagina que repetimos este proceso varias veces, manteniendo el mismo tamaño de la muestra y nivel de confianza, de tal forma que generemos muchos posibles intervalos. El nivel de confianza significa que **se espera que aproximadamente un** .  **de todos los intervalos contengan el verdadero valor de la media de la población**.

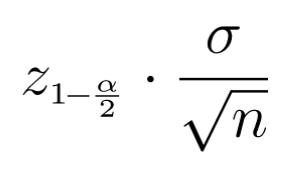
Lo anterior se ilustra en el siguiente gráfico que representa 100 intervalos de muestras de tamaño y con nivel de confianza del 95% cada uno:



Como se puede observar, de los 100 intervalos de confianza, 94 de ellos contienen la media . Si consideramos muchos más intervalos de confianza, construidos a partir de muestras de un tamaño razonable, deberíamos esperar que alrededor del 95% de ellos contenga a la media poblacional.

**ERROR DE ESTIMACIÓN**

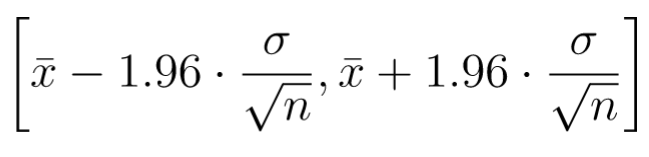
Esta expresión, utilizada para evaluar cuándo el valor de la media de la población pertenecía o no al intervalo de confianza, calculado a partir de una muestra aleatoria, se conoce como **error de estimación**, y se define como:

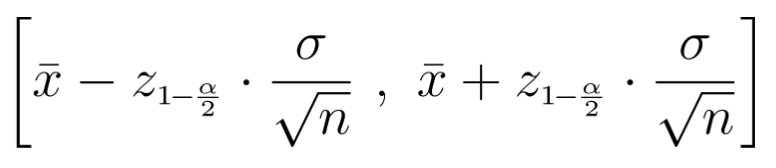
Notemos que el error de estimación corresponde a la mitad de la amplitud del intervalo de confianza.  


**SÍNTESIS**

* Imaginemos que queremos encontrar un intervalo que contenga el valor real de con una probabilidad del 95%. Esta probabilidad se suele denominar nivel de confianza, y puede ser distinta a 95%. Al intervalo que queremos encontrar lo denominamos intervalo de confianza.  
    
  Para determinar un intervalo de confianza para la media de la población hacemos dos supuestos:
* La distribución de la población sigue una distribución aproximadamente normal.
* El valor de la desviación estándar de la población es conocido.
* Considerando la distribución del promedio muestral , un intervalo con un 95% de confianza nos indica que la probabilidad de que el verdadero valor del parámetro se encuentre contenido dentro del intervalo es igual a 0.95.
* Para estimar los valores de y , se debe estandarizar la variable . Este desarrollo nos permite obtener un intervalo al 95% de confianza, es decir, la probabilidad de que el siguiente intervalo contenga al verdadero valor de es del 95%.



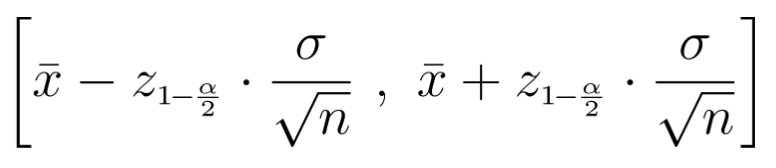
* Es importante aclarar que el intervalo obtenido es aleatorio, pues se genera a partir de la media muestral. El promedio muestral observado, denotado por corresponde a una realización de la variable aleatoria . Si reemplazamos por el promedio muestral observado , el intervalo fijo resultante se llama **intervalo de confianza de 95% para** . Este intervalo puede expresarse como:  
    
  
* Cuando estimamos intervalos de confianza, es usual querer estimarlos para distintos niveles de confianza. Por convención, el nivel de confianza se denota por , y expresa la probabilidad de que el intervalo determinado contenga al valor real del parámetro estimado.  
    
  Un intervalo de confianza de para una media poblacional , se puede expresar de la siguiente manera:



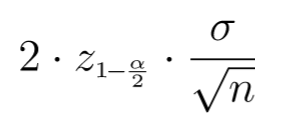
donde corresponde al promedio muestral observado, al tamaño de la muestra extraída desde la población, a la desviación estándar de la población y es el valor de asociado a un nivel de confianza .



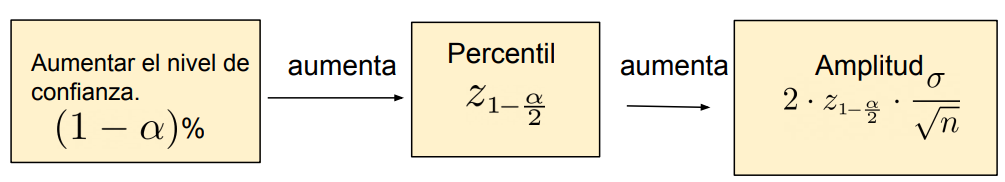
Considere el siguiente intervalo de confianza de para una media poblacional , con conocido:

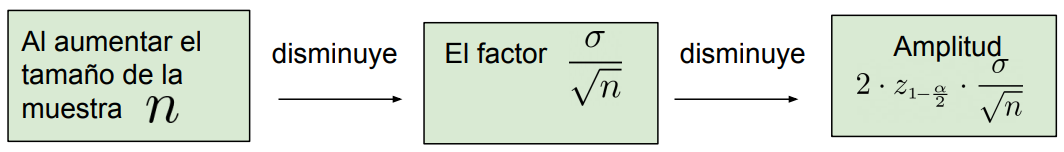


* La amplitud de este intervalo de confianza es igual a:



* Si aumentamos el nivel de confianza, manteniendo todos los demás términos constantes, la amplitud del intervalo de confianza (IC) aumenta. Esto se ilustra en la siguiente tabla:



* Si aumentamos el tamaño de la muestra, manteniendo todos los demás términos constantes, la amplitud del intervalo de confianza (IC) disminuye.   
    
  
* Considera una población donde se ha observado la variable de interés con media desconocida. Si seleccionamos un gran número de muestras de tamaño de esta población y a cada uno le calculamos su intervalo con un nivel de confianza , entonces:
* El nivel de confianza significa que se espera que aproximadamente un de todos los intervalos de confianza contenga el verdadero valor de la media de la población.
* Un valor pertenecerá a un intervalo de confianza si es que se mantiene a una distancia máxima  alrededor del promedio muestral. Esta distancia se denomina error de estimación.